Peer-review report of

Vu, D.-H., & Bierwiaczonek, K. (2025). Comparing bivariate and multivariate approaches to testing individual-level interaction effects in meta-analyses: The case of the integration hypothesis. *advances.in/psychology, 2*, e919144. https://doi.org/10.56296/aip00038

Round 1

Dear Drs Vu and Bierwiaczonek

Thank you for submitting your manuscript, **Examining the Interactive Nature of the Integration Hypothesis: A Comparison of Bivariate and Multivariate Meta-Analytical Approaches**, to our special issue in advances in/psychology.

Your manuscript was sent to two reviewers. These reviewers are both very familiar with the field of acculturation/intercultural-relations and meta-analyses. We have received comments from two reviewers who carefully evaluated your work. They both point to some strengths in the paper as well as highlight some areas that they think require further clarification and/or improvement.

I have read the paper and agree with the reviewers' comments. While I do not find it necessary to reiterate all their feedback, I will summarize some of their assessments, emphasizing the strengths of your study, points of agreement among the reviewers, and key unique concerns raised by each.

Strengths of the Paper

Both reviewers acknowledge the innovation and relevance of your study in the field of acculturation research. Reviewer 1 praises your contribution in applying a rigorous multivariate approach to estimating interaction effects in a meta-analysis, an improvement over previous studies that relied on bivariate methods. Your work is considered an important advancement for understanding the interplay between mainstream and heritage cultural engagement. Reviewer 2 recognizes the value of your study in re-analyzing two datasets using MASEM approaches and comparing them with other methodologies to assess empirical support for the interaction hypothesis.

Shared Reviewer Concerns

Both reviewers identify several overlapping concerns:

The availability of data: While your code is provided, both reviewers note that the referenced dataset is not accessible via the OSF repository, making replication of your analyses impossible.

The clarity and rationale of Study 3 (the simulation study): Both reviewers find it unclear why the simulation study was included and question its methodological setup

and purpose. They suggest either providing a stronger justification or removing it from the manuscript.

The need for a stronger discussion of prior relevant work: Reviewer 1 suggests revisiting the findings of Berry et al. (2022) in your General Discussion to provide a clearer contrast with your results.

Insufficient methodological details: Reviewer 2 points out that essential information about measures, inclusion/exclusion criteria, and sample details are not clearly presented, making it difficult for readers to interpret the study without referring to other sources.

Unique Comments from Each Reviewer

Reviewer 1:

Raises concerns about referencing decisions in the Introduction and suggests refining the discussion of past literature.

Questions the decision to ignore cross-lagged associations in the analysis, given that these could offer a stronger test of the hypothesis.

Highlights the importance of explaining how data was obtained from the original studies and whether additional steps were taken to ensure completeness.

Reviewer 2:

Suggests an alternative approach: Instead of using multiple potentially less suited methodological approaches, Reviewer 2 advocates pooling all available data into a single moderated regression analysis to test the interaction hypothesis more convincingly. Given that you have access to individual level data, the choice of a meta-analytic approach indeed becomes less clear.

Notes that the nesting of different ethnicities and host countries should be accounted for using a cross-classified multilevel analysis if raw data is available.

Raises concerns about the clarity of figures, particularly the difficulty in interpreting labels and colors in visual representations of results.

Decision and Next Steps

Based on the reviewers' feedback, I invite you to submit a major revision of your manuscript. In your revised submission, in addition to addressing each of the Reviewer’s comments, please ensure that you:

Provide access to the dataset referenced in your analysis or clarify any limitations regarding data availability.

Offer a stronger rationale for Study 3 or consider removing it.

Improve methodological transparency by detailing measures, sample characteristics, and criteria for study inclusion.

Address concerns about the choice of analytical methods, considering whether a pooled analysis might provide a more robust test of the interaction hypothesis.

Improve clarity in figures and graphs to enhance interpretability.

Engage more thoroughly with prior work, particularly Berry et al. (2022), to contextualize your findings.

I appreciate the effort you have invested in this research and believe that addressing these concerns will significantly strengthen your manuscript. I look forward to your resubmission and your response to the reviewers' comments within four weeks. Do not hesitate to reach out if you require more time.

Best regards,

David L. Sam

Reviewer 1:

The authors present an innovative new analysis of two meta-analytic datasets to test the hypothesis that acculturating individuals will fare better (in terms socio-cultural competence and psychological health) if they engage with both the mainstream and heritage cultures (*integration strategy*) rather than only with the heritage culture (*separation strategy*) or the mainstream culture (*assimilation strategy*). The authors observe, correctly, that this is a hypothesis about an interaction effect between the two variables and that past meta-analyses have used flawed strategies to estimate the interaction effect across studies. The present research makes a crucial contribution to the acculturation literature by conducting the first meta-analysis to use the appropriate multivariate approach to estimating interaction effects. In the remainder of this review, I will comment on each section of the proposal and highlight concerns that I would want the authors to address in a potential revision.

Introduction

The authors begin with a concise and coherent introduction to the integration hypothesis and end with an accessible review of bivariate approaches used to estimate interaction effects in past meta-analyses and their flaws. I have two minor comments. First, I did not quite understand why the first to references were included—“[m]any theories and hypotheses in psychology assume an interaction (Aguinis, 2002; Aguinis et al., 2005)”—as they are not necessary to make that

statement and do not seem to provide evidence for that statement. Second, the authors write in a footnote that “the only exception was Berry et al. (2022), who applied a typological approach (i.e., measures tapping into integration directly) to meta-analyze studies from the MIRIPS project, finding a correlation of .15”. As those findings contrast with those of the present research, I would expect the authors to revisit this meta-analysis in their *General Discussion* and to explain how its method differed from those used in the present research.

Studies 1 and 2

The two studies reanalyze data from two previous meta-analyses for which the authors were able to obtain individual-level data.

I have three minor comments. First, I would like to know more about how the authors went about trying to get this data. Did they consult the original papers? Did they contact authors to obtain this data? If not, why not? A meta-analysis lives and dies by its completeness so I think this is worth explaining. Second, I was confused about what the adaptation outcome measured and how different outcomes were combined. As I understand, past meta-analysis considered those outcomes separately. Third, I did understand why the authors ignored cross-lagged associations when those represent a stronger test of their hypothesis. As it is possible to analyze those within a MASEM framework, I would have expected to see this analysis in the paper.

I have one major comment. As far as I can see, the data referenced in the R script ("data_transformed.xlsx") is not provided in the OSF repository. Therefore, I was not able to reproduce the analyses. I refer to the editor on how that complies with journal policies.

Simulation Study 3

This part of the paper is the one I see as the biggest obstacle to publication. From reading the manuscript, it is not clear what the objective of the simulations are and why they are included. It is not surprising that analyses of simulated datasets based on the features of the existing data (used in Studies 1 and 2) produce similar results. A legitimate use of simulations would have been to estimate the sample size required to find an interaction if it did indeed exist but this does not seem what the authors did here. I urge the authors to either better explain the purpose of the simulations or to omit them from the manuscript.

General Discussion

The General Discussion does what a discussion section is supposed to do and outlines the practical and theoretical implications of the research. As mentioned under Introduction, I wish the authors would discuss their research in relation to the Berry et al. (2022) meta-analysis.

I hope the authors will find my comments useful for improving the present manuscript. I wish them all the best for their current and future research.

Reviewer 2:

The paper reports re-analyses of two different data sets using MASEM approaches and comparing them with four other approaches that have been reported in the literature but have problems. These are the midpoint split approach, the summative approach, the multiplative approach and the Euclidean distance approach. A simulation is reported in study 3. The stated objective of the paper is to address the question of the empirical support of the interaction hypothesis (see page 4 and page 7). One re-analyses with a two-stage MASEM yielded a non-significant interaction, the second re-analysis reported a significant interaction. The colours and labels for the plot are somewhat difficult to decipher. My interpretation of the graph is that when heritage culture orientation is high, the mainstream culture orientation is relatively less important. The greatest difference is observed at the mainstream orientation lower end - here individuals with high heritage culture seem to show better outcomes. It may have been useful to also plot the graph and run the simple effects with mainstream culture orientation as a moderator. There are reports of results for the alternative four approaches, but it is not clear to me what the value of reporting these results is given the conceptual and empirical problems outlined in the introduction.

My main observations are as follows:

a) In my casual reading of the literature, the interaction hypothesis is conceptually appealing but empirically weak. I am wondering why it is useful or beneficial to use four different approaches that are known to be weak. It would seem to be better in my opinion to pool all the available data into one single analysis and use appropriate methods to adquately test the interaction hypothesis.

b) The two re-analyses employed a MASEM design. If I understand correctly, the authors had access to the raw data (unfortunately, the data was not available on the OSF, therefore, I was not able to check or replicate the analyses). If you have access to the raw data, the best option would be to run a moderated regression, pooling across the two data sets. You can include methodological or sample artefacts as control variables or moderators. You might be able to control for dependency in the data (e.g., ethnicities and cultures sampled in different studies)

directly via including random intercepts. This would be a strong and convincing test of the interaction hypothesis.

c) I did not understand the rationale for using MASEM. If you have access to the raw data, it would be best to directly re-analyse the data at the individual level (with inclusion of appropriate random effect structures). In my reading of the code, none of the advantages of MASEM to account for measurement unreliability etc are being used. It was unclear to me why the MASEM approach was preferable to a re-analysis with (multilevel) regression, assuming that you have access to the raw data (as was mentioned in a number of places in the manuscript).

d) I did not understand the rationale and the set up of the simulation study. It would be important to clearly explain and justify the aims of the simulation study, discuss why certain parameters were modulated and the rationale for why these parameters are important or relevant and finally, a careful presentation of the main results. Why is the number of studies relevant? When I read the early parts of the paper I was expecting a simulation of parameters that would identify crucial parameters of the interaction parameters (e.g., size of the interaction, correlation of the predictors/moderators, reliability of the predictors/moderators, sample sizes of the original studies and the like) to see under what conditions it is possible to detect an interaction in line with the hypotheses. I was not certain that the current set up was convincing. Why number of studies? What parameters did you simulate in each of the studies? How many studies were at the lower and higher end of the correlation bound?

e) I was not convinced of the value of presenting the four different approaches. I think it is valid to mention them, to highlight their problems and then proceed with an analysis that is empirically sound to test what is at stake (the integration hypothesis).

f) One issue that I think needs consideration is the nesting of different ethnicities within and across countries. This dependency is not properly discussed or acknowledged. If you have access to the raw data (as was mentioned in the ms at various points), then I would suggest a cross-classified multilevel analysis, with two different random effects for ethnicity and host country.

g) as noted above - the code is available, but the data is not. It was therefore not possible to replicate the analyses.

h) There was very limited information on the measures, samples and other relevant information that are normally part of any meta-analysis. What measures were studied, how were they combined, what samples, what inclusion and exclusion criteria, etc. It would be necessary to provide sufficient information to a reader to understand the current paper without needing to read other sources to be able to interpret the results.

In summary, the rationale and main point of the paper was unclear to me. Study 3 is poorly explained and it is not clear what the main point of that analysis might be. The main contribution of this paper could be a proper re-analysis of the existing datasets to examine whether the interaction hypothesis is empirically supported or not (because this seems to be the main objective if I understand the arguments on

pages 4 and 7 correctly). If this is indeed the main objective, I would recommend pooling the two studies into one proper re-analysis. The interaction effect could then be directly tested. Methods, measurement and sampling artefacts could be controlled for or included as moderators.

Round 2:

Dear Authors,

Thank you for submitting your revised manuscript, "**Comparing Bivariate and Multivariate Approaches to Testing Individual-Level Interaction Effects in Meta-Analyses: The Case of the Integration Hypothesis**" (formerly titled "*Examining the Interactive Nature of the Integration Hypothesis: A Comparison of Bivariate and Multivariate Meta-Analytical Approaches*") to our special issue in Advances in Psychology.

I have now received feedback from the original reviewers. While Reviewer 1 noted some improvements and appreciates the effort you've put into the revision, Reviewer 2 remains more critical. Both reviewers, however, recognize the importance of your topic—exploring methods for estimating interaction effects in meta-analyses, especially within acculturation research.

After reading the manuscript myself, I've summarized the key areas of agreement and disagreement, as well as suggestions for moving forward.

Key Points of Agreement

1. Concerns About Simulation Study 3
    - Reviewer 1 feels this study reiterates well-known statistical principles (e.g., power and sample size) without offering new insights.
    - Reviewer 2 also questions its purpose and contribution, suggesting it doesn't significantly enhance understanding of interaction effects.
2. Limitations of the Four Analytical Methods
    - Reviewer 2 agrees that the four methods (midpoint split, summative scores, multiplicative scores, and Euclidean distance) have notable shortcomings.

Points of Disagreement

- Reviewer 1 is generally supportive of the revised manuscript and would recommend publication if Simulation Study 3 is removed.
- Reviewer 2, however, remains unconvinced about the paper's overall contribution.

Next Steps

Despite the mixed reviews, I believe your paper has potential to make a valuable contribution to the special issue, because the methods you compare are indeed still vastly, and often uncritically, used in the field, including in influential meta-analyses. There, thus, is value in the present contribution. I'd therefore like to invite you to revise and resubmit your manuscript. If you choose to proceed, please address the following:

Move Simulation Study 3 to the SOM

Both reviewers are experts in meta-analysis and repeatedly disagreed with the value of the simulation. As a compromise, please move it to the SOM, briefly summarize its result in the main text and refer readers to it.

In sum, the paper requires only minor revisions. I will kindly ask that you submit your revised manuscript by June 30.

Thank you again for your thoughtful work and your contribution to this important area of research. I look forward to seeing your next version soon and plan to make the final decision then.

Sincerely

David L. Sam

Reviewer 1:

I thank the authors for engaging with my comments and suggestions. I am satisfied with the authors' revisions and responses concerning most of the concerns raised in my review. Unfortunately, I still do not see the value of Simulation Study 3.

Simulation Study 3

*With parameters varying this widely, the range of average effects that can be found in meta-analyses is likely to be extreme regardless of the method used, and without a larger dataset **(or a simulation)** reproducing this variability, it is impossible to know to what extent diverging results are attributable to methods rather than error, and by consequence, how trustworthy the bivariate approaches used in previous meta-analyses really are. … (1) What range of interaction effect sizes can be obtained using the multivariate approach? (2) How often, if ever, do bivariate approaches produce similar effect sizes as the multivariate approach? (3) How likely are we to find a significant interaction effect and what parameters (e.g., number of studies, intercorrelations between variables) predict it?*

I agree with the author that between-study variability makes it difficult to interpret the results of meta-analyses and this is a particular issue when meta-analyses rely of relatively few effect sizes. While more data would help (when doesn't it?), I continue to disagree with the authors that their simulation study helps address this issue. As

the simulations are based on parameters obtained from Studies 1 and 2, it is not clear to me how they could possibly provide information that is not already contained in Studies 1 and 2. This becomes even clearer when reviewing the conclusions the authors draw from the simulations. For example, the authors conclude that (1) “the key factor affecting the likelihood of detecting a significant interaction seems to be the sample size” and (2) “the few smaller meta-analyses (k = 20) that do show significant effects tend to have relatively larger effect sizes”. In other words, they conclude that (1) statistical power is a function of sample size and (2) smaller studies with significant results will have larger effect sizes (winner’s curse)—both of which are inherent to the logic of null hypothesis significance testing. In view of this, I recommend that the authors remove Simulation Study 3.

Besides this issue, I am happy to recommend the study for publication as the discussion of different approaches to estimating interaction effects in meta-analyses seems like an interesting and worthwhile contribution to the literature.

Reviewer 1:

Thank you for your revisions. I do appreciate your responses, but I still struggle to see the objective of your paper. The four methods that you present are clearly problematic. I fail to see why a demonstration of their effects or even a simulation to see what the effects do look like is meaningful in any way.

1) Midpoint Split

This approach results in a loss of information, because dichotomizing continuous variables leads to substantial information loss, reducing variability and statistical power. It also suffers from the choice of arbitrary Thresholds (e.g., is the choice of midpoint or median theoretically meaningful? what is the sample effect?). The results are not comparable across samples or studies.

2) Summative scores

This approach presumes additivity which is antithetical to an interaction hypothesis. It is impossible to disentangle whether the outcome is driven by one orientation or the interplay between both.

3) Multiplicative scores

This is somewhat better but still flawed. While this approach attempts to capture an interaction, it does so outside the full context of a moderated regression model, where main effects and interaction effects are simultaneously modeled and interpreted. The interpretation of the combined term is ambiguous.

4) Euclidean Distance (ED)

I was not familiar with this approach. It seems to be more of a topological arrangement of the data. I am not sure how well it is suited for testing an interaction effect.

In summary, these methods do have obvious flaws and issues. Why do we need to give them more airway? It would make more sense to clearly demonstrate how to properly test the interaction effect. As you need to have access to the raw data, it seems to make most sense to test the interaction effect directly with a moderated multiple regression. There are sufficient complexities with this approach that need to be discussed and compared (e.g., the renewed question on centering, baseline effects) that this should be a much better use of the journal space and resources.

Acculturation research has been criticized for low rigor and quality. The current paper does not seem to help much. A clear summary of the criticisms together with a clear description of how to move forward, followed by a re-analysis of the existing data to address the core question about the interaction effect would have an impact. A simulation on parameter or model choices as well as the role of sample size could be useful. But first show what method is most appropriate and what the available data says about the theoretical question.